

# **EXHIBIT 31**

**SUPPLEMENTAL REPORT SUBMITTED IN THE MATTER OF**

**Amy Bartoletti, Chia Siu, Nadine Mentor,  
Lisa Conley, and Brittany Sharpton**

**v.**

**Citigroup Inc. and Citigroup Global Markets Inc.**

**Civil Action No. 10 Civ. 7820 (PJO)  
United States District Court  
Southern District of New York**

**Professor David E. Bloom**

**Harvard University**

**February 12, 2013**

I have reviewed “Expert Report of Mark R. Killingsworth, D.Phil.” dated December 11, 2012 (hereafter referred to as the Killingsworth Report). I have also reviewed “Expert Report: Criteria for Layoffs and Gender Disparities in Citigroup’s Public Finance Department: The case of Amy Bartoletti, Chia Siu, Nadine Mentor, Lisa Conley, and Brittany Sharpton v. Citigroup Inc. and Citigroup Global Markets Inc.” by Louise Marie Roth, dated December 10, 2012 (hereafter referred to as the Roth Report). The Killingsworth Report endeavors “to determine whether (and if so to what extent) there were sex-related disparities in the composition of terminations within the Public Finance Department (“PFD”) during ...[Citigroup’s November 2008] reduction-in-force (“RIF”)” (Killingsworth Report, page 1, item 3). The Roth Report endeavors “to assess whether the criteria used to select employees for termination in the Public Finance Department during the November 2008 RIF contributed to gender disparities in layoffs” (Roth Report, p. 1, item 1).

## I. Comments on the Killingsworth Report

### *A. The Killingsworth Report inappropriately treats the layoffs of the five plaintiffs as if they resulted from a single process involving all the employees in the Public Finance Department.*

As described in the Bloom Report dated December 11, 2012, and acknowledged by Dr. Roth in the Roth Report, it is clear from the deposition testimony in this case that the five plaintiffs were laid off as a result of multiple processes that operated at the level of the group and not at the level of the Public Finance Department. On page 8 (item 21) of the Roth Report, Dr. Roth explains that “Brownstein and Chin left it up to the group heads to select people for termination....” As noted in the Bloom Report (page 3), “Individuals in the same group, and in the same or similar positions, were compared against each other, but not against individuals in fundamentally dissimilar positions or in other groups.” This institutional reality is mirrored closely in the statistical analyses contained in the Bloom Report. However, it is ignored in the Killingsworth Report.

The Killingsworth Report effectively treats the layoffs of the five plaintiffs as if they resulted from a single process involving all the employees in the Public Finance Department. However, the reality is that the selection of the five plaintiffs for layoff resulted from multiple processes involving distinctly different groups, positions, decision-makers, and/or comparators. Indeed, roughly half of the 145 employees in Dr. Killingsworth’s probit analysis were not even situated within the four groups that employed the five plaintiffs at the time of RIF IV. Outcomes associated with these employees are not meaningful with respect to statistical inferences regarding the plaintiffs’ claims that “...during the course of a reduction-in-force that Citigroup undertook on or about November 21, 2008, Citigroup terminated their employment because of their sex” (Killingsworth Report, page 1, item 2). In particular, the Killingsworth analysis treats layoffs in Credit and Financial Products, Power, Central, Mid-Atlantic, Northeast, Southwest, and West as if they are relevant to layoff selections in Healthcare, Housing, Infrastructure, and Southeast, which they are not.

In addition, the Killingsworth Report treats layoff selections in each of the four plaintiff groups as if they are relevant to layoff selections in the other three plaintiff groups, which they are not. For example, the Killingsworth Report inappropriately treats layoffs in the Housing Group as if they are relevant to assessing gender disparities in layoff selections in the Healthcare Group, the Infrastructure Group, and the Southeast Group.

Statistical tests that pay careful attention to appropriate comparators for each of the five plaintiffs are the centerpiece of the Bloom Report. These tests find no statistically significant gender disparities in selection for layoff among individuals who were similarly situated to each plaintiff with respect to their job/position and group.

***B. The Killingsworth Report fails to examine data on, or otherwise account for, RIF Rounds I, II, and III.***

The statistical analyses in the Killingsworth Report refer exclusively to the reduction in force (RIF) that was associated with notice dates from November 21 through December 18, 2008. The Killingsworth Report does not mention or analyze any data related to the RIFs associated with notice dates from

- December 18, 2007 through January 22, 2008 (RIF I),
- March 18-19, 2008 (RIF II), or
- June 23 through July 7, 2008 (RIF III).

The outcomes of RIFs I through III provide relevant information about layoff decision-making because, in each of those prior rounds, plaintiffs Bartoletti, Conley, Mentor, Sharpton, and Siu were all at risk of layoff, they were compared to many of the same employees, and selections for layoff were made by predominantly the same decision-makers.<sup>1</sup>

Although it would be inappropriate for the reasons described in Section I.A. above, had Dr. Killingsworth performed the same analysis he does for RIF IV on RIFs I, II, and III, he would have obtained the results set forth in Panels A through C of Tables BSR1 and BSR2 below.

Table BSR1 corresponds to Table 1 in the Killingsworth Report. While Dr. Killingsworth finds a statistically significant “sex difference in termination rates” between males and females in RIF

---

<sup>1</sup> The one exception is the Housing Group, which was headed by Nicholas Fluehr until he was laid off as part of RIF III. Following Mr. Fluehr’s departure, Ms. Bartoletti and Mr. Michael Koessel were named Co-Heads of the Housing Group, although Mr. David Brownstein (who joined Francis Chin as Co-Head of the Public Finance Department in early 2008) appears to have selected employees for RIF III and RIF IV. Also, Nadine Mentor was at risk for RIF I as a Vice President but was at risk for RIFs II, III, and IV as a Director, indicating that her relevant comparators changed over the span of the four RIFs.

IV, differences in the termination rates of males and females are not statistically significant for RIF I, RIF II, or RIF III. Note that in RIF III, the termination rate for females is lower than that for males; in RIF II, the male and female termination rates are almost identical.

Table BSR1, Panel A  
 Sex difference in termination rates,  
 Citigroup Public Finance Department, RIF I  
 For Comparison with Table 1 in Killingsworth Report

<u>employee group</u>	<u>total number</u>	<u>terminated</u> <u>number</u>	<u>% of total</u>
female	48	8	16.7%
male	153	14	9.2%
female-male difference:			7.5%
<u>statistical significance:</u>			
number of "standard errors"			1.455
p-value ("chance probability")			0.1456

Table BSR1, Panel B  
 Sex difference in termination rates,  
 Citigroup Public Finance Department, RIF II  
 For Comparison with Table 1 in Killingsworth Report

<u>employee group</u>	<u>total number</u>	<u>terminated</u> <u>number</u>	<u>% of total</u>
female	39	1	2.6%
male	139	3	2.2%
female-male difference:			0.4%
<u>statistical significance:</u>			
number of "standard errors"			0.1511
p-value ("chance probability")			0.8799

Table BSR1, Panel C  
 Sex difference in termination rates,  
 Citigroup Public Finance Department, RIF III  
 For Comparison with Table 1 in Killingsworth Report

<u>employee group</u>	<u>total number</u>	<u>terminated</u> <u>number</u>	<u>% of total</u>
female	42	2	4.8%
male	136	13	9.6%
female-male difference:			-4.8%
<u>statistical significance:</u>			
number of "standard errors"		0.9782	
p-value ("chance probability")		0.3280	

Table BSR2 corresponds to Table 2 in the Killingsworth Report. While Dr. Killingsworth finds a statistically significant female-male difference in the probability of termination for RIF IV, the corresponding differences are not statistically significant for RIF I, RIF II, or RIF III. Note that in RIF II and in RIF III, the point estimate of the female-male difference in probability of termination is negative (i.e., the conditional RIF rate is lower for females than for males).

Table BSR2, Panel A  
 Female-male difference in probability of termination  
 for Citigroup Public Finance Department employees, RIF I  
 (summary of probit analysis)  
 For Comparison with Table 2 in Killingsworth Report

	probit coefficient (z statistic) [p-value]
female-male difference in probability of termination	0.268 (0.83) [0.409]

Table BSR2, Panel B  
 Female-male difference in probability of termination  
 for Citigroup Public Finance Department employees, RIF II  
 (summary of probit analysis)  
 For Comparison with Table 2 in Killingsworth Report

	probit coefficient (z statistic) [p-value]
female-male difference in probability of termination	-0.077 (-0.09) [0.925]

Table BSR2, Panel C  
 Female-male difference in probability of termination  
 for Citigroup Public Finance Department employees, RIF III  
 (summary of probit analysis)  
 For Comparison with Table 2 in Killingsworth Report

	probit coefficient (z statistic) [p-value]
female-male difference in probability of termination	-0.400 (-0.79) [0.427]

***C. Other Comments on the Killingsworth Report.***

- a. Dr. Killingsworth uses the word “department” to refer to “groups” within the overall Public Finance Department. For example, see page 2, footnote 1 in the Killingsworth Report.
- b. Item 5 should read “less than .7 times in 100” or “less than 7 times in 1000” rather than “less than 7 times in 100.”
- c. There is a discrepancy between Item 10 and Table 3. Item 10 reads “the actual percentage terminated was 34.3 percent” whereas the corresponding figure in Table 3 is 14.3 percent.

## II. Comments on the statistical analyses in the Roth Report

The statistical analyses in the Roth Report have several major flaws. First, Dr. Roth uses statistical tests that are not appropriate to the “small samples” she examines. Second, Dr. Roth is inappropriately selective in her reporting of statistical results, and fails to report findings pertaining to RIF Rounds I, II, and III, even though she received and examined data for those RIF rounds. Third, Dr. Roth offers no analyses that compare individuals who are similarly situated to the five plaintiffs with respect to their group and position.

### *A. Group-level comparisons.*

In Figure 1 on page 7 of the Roth Report (mistakenly referred to as “Figure 2” in item 18 on page 6), Dr. Roth displays a bar chart showing the male and female layoff rates in RIF IV for the four plaintiff groups: Health Care, Housing, Infrastructure, and Southeast. Dr. Roth notes that “the number of cases in each subgroup prohibited meaningful statistical analysis by subgroup for all subgroups within the PFD” (Roth page 6, point 17). Dr. Roth goes on to report statistically significant disparities for two of the four groups: Housing (7 at-risk employees and 4 layoffs in RIF IV) and Infrastructure (18 at-risk employees and 5 layoffs in RIF IV). Dr. Roth does not indicate the particular statistical test she used to reach this conclusion, but the p-values she provides on page 6 of the Roth Report (“p<.01” for Infrastructure and “p<.05” for Housing) are consistent with a chi-square test. The chi-square test is an asymptotic test and is not appropriate given the small numbers of female and male at-risk employees and of female and male layoffs. By contrast, the Fisher’s Exact Test is appropriate to these sample sizes, as noted in the Bloom Report. The Fisher’s Exact Test indicates no statistically significant difference between male and female layoff rates within the Housing Group, Southeast Group, or Healthcare Group in RIF IV.

Healthcare is, in fact, the largest of all of the Public Finance Department groups, with 36 at-risk employees and 4 layoffs in RIF IV, and Dr. Roth’s programs show that she did compare male and female layoff rates within this group (file “descr6.pdf,” page 70; note that Dr. Roth only identifies 34 at-risk employees in the Healthcare group, which deviates from the 36 at-risk employees identified in the back-up files produced by Dr. Killingsworth and in the Bloom Report dated December 11, 2012).<sup>2</sup> Although inappropriate for samples of this size, the chi-square test indicates a statistically insignificant difference between male and female layoff rates within the Healthcare Group. Dr. Roth’s programs also show that she compared male and female layoff rates within the Southeast Group (file “descr6.pdf,” page 71), with 13 at-risk employees and 2 layoffs in RIF IV. The chi-square test indicates a statistically insignificant

---

<sup>2</sup> There are additional discrepancies between the group counts in the Killingsworth and Roth analyses. These involve the following groups: Administration, Airports, Credit and Financial Products, and Northeast, in addition to Healthcare. Note that the Killingsworth counts match those provided in the Bloom Report.

difference between male and female layoff rates within the Southeast Group. Dr. Roth does not report these results.

The results in Figure 1 of the Roth Report refer only to outcomes related to RIF IV, though Dr. Roth refers to RIFs beginning in December 2007 on page 4 (item 11) and again on page 5 (item 14). The page 4 reference states explicitly that Dr. Roth received personnel data corresponding to all four RIF rounds: "Citi turned over a spreadsheet of data on the personnel in its Municipal Securities Division before and after its RIF from December 2007 to November 2008." If Dr. Roth had repeated the Figure 1 analysis for the earlier RIF rounds, she would have found no significant differences in female vs. male layoff rates for any of the four plaintiff groups, although in some cases the test cannot be constructed. The absence of statistically significant disparities also holds using the Fisher's Exact Test for analyses in which it can be constructed.

Dr. Roth's Figure 1 analysis is also flawed in that it takes no account of the positions held by the plaintiffs or their comparators. Group heads evaluated employees relative to those who were similarly situated in terms of their position in the career progression from Analyst to Managing Director. Statistical results that ignore position therefore cannot be used to evaluate the plaintiffs' claims because they involve comparisons of employees who were not similarly situated. By contrast, statistical tests that pay careful attention to appropriate comparators for each of the five plaintiffs are central to the Bloom Report. These tests find no statistically significant gender disparities in selection for layoff among individuals who were similarly situated to each plaintiff with respect to their job/position and group. The findings hold for RIF IV, and also for RIF I, RIF II, and RIF III.

#### ***B. Logistic Regression for Directors.***

On page 6 (item 16) of the Roth Report, Dr. Roth reports the result of a logistic regression that compares the female vs. male likelihood of layoff in RIF IV for Directors. Had Dr. Roth applied the same test to Directors for RIF III, she would have found no significant gender disparity ( $p=.89$ ). The test cannot be applied to Directors in RIF I or RIF II because there were no layoffs among female Directors in RIF I or RIF II. This is in contrast to 5 out of 59 male Directors laid off in RIF I and 1 out of 56 male Directors laid off in RIF II.

Moreover, Dr. Roth's logistic regression for Directors does not differentiate between Directors in different groups within the Public Finance Department, even though Dr. Roth's summary of the deposition testimony acknowledges that layoffs in the Public Finance Department resulted from multiple group-level processes. By contrast, statistical tests that pay careful attention to appropriate comparators for each of the five plaintiffs are at the heart of the Bloom Report. These tests find no statistically significant gender disparities in selection for layoff among individuals who were similarly situated to each plaintiff with respect to their job/position and group. The findings hold for RIF IV, and also for RIF I, RIF II, and RIF III.

Dr. Roth applies her page 6 logistic regression selectively, only to Directors. Had Dr. Roth estimated her logistic regression using data for Analysts (referring to plaintiffs Sharpton and Siu), she would have found no significant disparity in RIF I ( $p=.68$ ) or in RIF IV ( $p=.16$ ). The test cannot be applied to Analysts in RIF II because there were no layoffs among Analysts in that round. The test cannot be applied to Analysts in RIF III because there were no layoffs among female Analysts in that round. By comparison, one of 18 male Analysts was laid off in RIF III. Had Dr. Roth estimated her logistic regression using RIF I data for the subsample of employees with the title of Vice President (referring to plaintiff Mentor in RIF I), she would have found no significant gender disparity ( $p=.72$ ).

***C. Results for the entire Public Finance Department.***

On page 5 of the Roth Report (item 14), Dr. Roth states that “During the 12-month period from December 2007 to late November 2008, the PFD laid off 70 employees. The layoffs disproportionately targeted female employees.” Dr. Roth then discusses the results of a logistic regression that compares the female vs. male likelihood of layoff using data for RIF IV only, and from the entire Public Finance Department (page 5, item 14). Had Dr. Roth applied the same logistic regression analysis to RIF I, RIF II, and RIF III, she would have found no significant gender disparities for any of those RIF rounds ( $p=.15$  for RIF I,  $p=.88$  for RIF II, and  $p=.34$  for RIF III). Indeed, the Roth Report and corresponding statistical production clearly establish that Dr. Roth received and reviewed the data needed to conduct such regression analyses.

Table 1 of the Roth Report (“Gender Differences in Layoffs, Nov. 2008”) offers fundamentally the same result as the logistic regression discussed by Dr. Roth on page 5. In Table 1, Dr. Roth performs a chi-square test of statistical significance, with the p-value in the table notes. If Dr. Roth had performed her Table 1 analysis for RIFs I, II, and III, she would have obtained the results set forth in Panels A through C of Table BSR3 below. Differences in layoffs of males vs. females are not statistically significant in any of RIFs I, II, or III. Note that in RIF III, the layoff rate for females is lower than that for males and in RIF II, the male and female layoff rates are almost identical.

**Table BSR3, Panel A: Gender Differences in Layoffs, RIF I**  
 For comparison with Table 1 in Roth Report

	Male	Female	Total
Remained after RIF I	139 (90.9%)	40 (83.3%)	179 (89.1%)
Laid off in RIF I	14 (9.2%)	8 (16.7%)	22 (11.0%)
Total	153	48	201
Pearson Chi-Square	2.12		

Note: Percentages are of men and of women in the PFD  
 p-value for Pearson Chi-Square = 0.15 (not statistically significant)

**Table BSR3, Panel B: Gender Differences in Layoffs, RIF II**  
 For comparison with Table 1 in Roth Report

	Male	Female	Total
Remained after RIF II	136 (97.8%)	38 (97.4%)	174 (97.8%)
Laid off in RIF II	3 (2.2%)	1 (2.6%)	4 (2.3%)
Total	139	39	178
Pearson Chi-Square	0.02		

Note: Percentages are of men and of women in the PFD  
 p-value for Pearson Chi-Square = 0.88 (not statistically significant)

**Table BSR3, Panel C: Gender Differences in Layoffs, RIF III**  
 For comparison with Table 1 in Roth Report

	Male	Female	Total
Remained after RIF III	123 (90.4%)	40 (95.2%)	163 (91.6%)
Laid off in RIF III	13 (9.6%)	2 (4.8%)	15 (8.4%)
Total	136	42	178
Pearson Chi-Square	0.96		

Note: Percentages are of men and of women in the PFD  
 p-value for Pearson Chi-Square = 0.33 (not statistically significant)

Moreover, neither the page 5 logistic regression result nor the Table 1 chi-square result focuses on employees who were similarly situated to the plaintiffs and their comparators. Group heads evaluated employees relative to those who were similarly situated in terms of their position in the career progression from Analyst to Managing Director. Statistical results that ignore job position therefore cannot be used to evaluate the plaintiffs' claims because they involve comparisons of employees who were not similarly situated. By contrast, and as noted earlier, statistical tests that pay careful attention to appropriate comparators for each of the five plaintiffs are central to the Bloom Report. These tests find no statistically significant gender disparities in selection for layoff among individuals who were similarly situated to each plaintiff with respect to their job/position and group. The findings hold for RIF IV, and also for RIF I, RIF II, and RIF III.



---

David E. Bloom  
February 12, 2013